CSE 151 Machine Learning
Instructor: Kamalika Chaudhuri
Announcements

**Final:** June 13, 7-9pm (2 hour final)

**Material Covered:** Everything! (Linear algebra, Probability review to Gaussian Mixture Models)

Review session will be announced

Homework 4 is up, due Mon Jun 4 in class
Projecting Data

How to find the subspace $S$ to project to?

One option: PCA subspace
**Eigenvectors and Eigenvalues**

Given a \( n \times n \) matrix \( A \),

- \( c \): eigenvalue of \( A \)
- \( x \): the corresponding eigenvector if:

\[
Ax = cx \quad x \neq 0
\]
Symmetric $A$

1. All $c_i$ are real

2. All $x_i$ are orthonormal

$$A \ X = X \ C$$

$$A = X \ C \ X^T \quad \text{(for symmetric $A$)}$$

since $X^T = X^{-1}$ for orthonormal matrices

$X = \text{an orthonormal basis of } \mathbb{R}^n$
Properties of Symmetric A

\[
\max_{x \in \mathbb{R}^n} x^\top A x
\]

Such that: \[\|x\| = 1\]

**Solution:** \(x = x_1\) (eigenvector for the largest eigenvalue)
maximum value of \(x^\top A x\) is the largest eigenvalue

**Question:** If \(x\) is an eigenvector with eigenvalue \(c\), what is \(x^\top A x\)?
Properties of Symmetric $A$

\[
\max_{X \in \mathbb{R}^{n \times k}} \text{trace}(X^\top AX)
\]

Such that columns of $X$ are orthonormal

**Solution:** $X = [x_1 \ x_2 \ldots \ x_k]$

(eigenvectors for the $k$ largest eigenvalues)

maximum value of $\text{tr}(X^\top AX)$ is sum of top $k$ eigenvalues
In Class Problem

Let $A$ be a symmetric $d \times d$ matrix

$X = [x_1, \ldots, x_k]$, where $x_i$ is the $i$-th eigenvector of $A$

Each $x_i$ has norm 1

Show that:

(a) Entry $(i, i)$ of $X^TAX$ is $c_i$, the $i$-th eigenvalue of $A$

(b) Entry $(i, j)$ of $X^TAX$ is 0, when $i$ is not equal to $j$
Suppose $x_1, \ldots, x_n$ are data points in $\mathbb{R}^d$

Covariance matrix of $x_1, \ldots, x_n$:

$$\Sigma = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^\top$$

where:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$
Covariance Matrix

Covariance matrix of \(x_1, \ldots, x_n\):

\[
\Sigma = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^\top
\]

where:

\[
\mu = \frac{1}{n} \sum_{i=1}^{n} x_i
\]

**Properties:**

1. Entry at row \(i\), column \(j\) is covariance of coord. \(i\) and coord. \(j\)
2. Symmetric
3. Positive semi-definite (0 or positive eigenvalues)
How to find a good subspace?

Data lies on a line

Direction 1

Direction 2

Direction 1 projections

Direction 2 projections
How to find a good subspace?

Data lies almost on a line
Find the direction of **maximum variance**
Max Variance Direction

Suppose data in $\mathbb{R}^d$ is centered so that the mean is 0

Variance along a direction $u$ is:

$$
\frac{1}{n} \sum_{i=1}^{n} (x_i, u)^2 = \frac{1}{n} \sum_{i=1}^{n} (u^T x_i)^2 = \frac{1}{n} \sum_{i=1}^{n} u^T x_i x_i^T u = \frac{1}{n} u^T \left( \sum_i x_i x_i^T \right) u
$$

Recall that: $A = \sum_i x_i x_i^T$ is symmetric

Finding the max. variance direction is equivalent to finding the top eigenvector of $A$. Similar fact can be shown for the k-dimensional subspace of max variance.
Dimension Reduction Algorithm

Given data matrix $X_{n \times d}$

1. Compute mean $m$ of the data vectors

2. Compute covariance matrix:

$$
\Sigma = \frac{1}{n} \sum_{i=1}^{n} (x_i - m)(x_i - m)\top
$$

3. Compute $U_{k \times d}$, the top $k$ eigenvectors of $\Sigma$

4. Project:

$$
P(X) = XU\top
$$